

A SIMPLE MODEL OF TREES FOR UNICELLULAR MAPS

GUILLAUME CHAPUY, VALENTIN FÉRAY AND ÉRIC FUSY

ABSTRACT. We consider unicellular maps, or polygon gluings, of fixed genus. A few years ago the first author gave a recursive bijection transforming unicellular maps into trees, explaining the presence of Catalan numbers in counting formulas for these objects. In this paper, we give another bijection that explicitly describes the “recursive part” of the first bijection. As a result we obtain a very simple description of unicellular maps as pairs made by a plane tree and a permutation-like structure. All the previously known formulas follow as an immediate corollary or easy exercise, thus giving a bijective proof for each of them, in a unified way. For some of these formulas, this is the first bijective proof, e.g. the Harer-Zagier recurrence formula, or the Lehman-Walsh/Goupil-Schaeffer formulas. Thanks to previous work of the second author this also leads us to a new expression for Stanley character polynomials, which evaluate irreducible characters of the symmetric group.

1. INTRODUCTION

A unicellular map is a connected graph embedded in a surface in such a way that the complement of the graph is a topological disk. These objects have appeared frequently in combinatorics in the last forty years, in relation with the general theory of map enumeration, but also with the representation theory of the symmetric group, the study of permutation factorizations, or the study of moduli spaces of curves. All these connections have turned the enumeration of unicellular maps into an important research field (for the many connections with other areas, see [12] and references therein; for an overview of the results see the introductions of the papers [5, 1]). The counting formulas for unicellular maps that appear in the literature can be roughly separated into two families.

The first family deals with *colored* maps (maps endowed with an application from its vertex set to a set of q colors). This implies “summation” enumeration formulas (see [10, 18, 14] or paragraph 3.4 below). These formulas are often elegant, and different combinatorial proofs for them have been given in the past few years [13, 8, 18, 14, 1]. The issue is that some important topological information, such as the genus of the surface, is not apparent in these constructions.

Formulas of the second family keep track explicitly of the genus of the surface. This includes inductive relations (like the Harer-Zagier recurrence formula [10]) or explicit (but quite involved) closed forms (Lehman-Walsh [21] and Goupil-Schaeffer [9] formulas). From a combinatorial point of view, these formulas are harder to understand. A step in this direction was done by the first author in [5] (this construction is explained in subsection 2.2), which led to new induction relations and to new

Key words and phrases. one-face map, Stanley character polynomial, bijection, Harer-Zagier formula, Rémy’s bijection.

First and third author partially supported by the ERC grant StG 208471 – ExploreMaps.

Second author partially support by ANR project PSYCO.

formulas. However the link with other formulas in the second family remained mysterious, and [5] left open the problem of finding combinatorial proofs of these formulas.

The goal of this paper is to present a new bijection between unicellular maps and surprisingly simple objects which we call *C-decorated trees* (these are merely plane trees equipped with a certain kind of permutation on their vertices). This bijection is based on the previous work of the first author [5]: we explicitly describe the “recursive part” appearing in this work. As a consequence, not only can we reprove all the aforementioned formulas in a bijective way, thus giving the first bijective proof for several of them, but we do that in unified way. Indeed, C-decorated trees are so simple combinatorial objects that all formulas follow from our bijection as an immediate corollary or easy exercise.

Another interesting application of this bijection is a new explicit way of computing the so-called Stanley character polynomials, which are nothing but the evaluation of irreducible characters of the symmetric groups, properly normalized and parametrized. Indeed, in a previous work [6], the second author expressed these polynomials as a generating function of (properly weighted) unicellular maps. Although we do not obtain a “closed form” expression (there is no reason to believe that such a form exists!), we express Stanley character polynomials as the result of a term-substitution in free cumulants, which are another meaningful quantity in representation theory of symmetric groups.

2. THE MAIN BIJECTION

2.1. Unicellular maps and C-decorated trees. A *map* M of genus $g \geq 0$ is a connected graph G embedded on a closed compact oriented surface S of genus g , such that $S \setminus G$ is a collection of topological disks, which are called the *faces* of M . Loops and multiple edges are allowed. The graph G is called the *underlying graph* of M and S its *underlying surface*. Two maps that differ only by an oriented homeomorphism between the underlying surfaces are considered the same. A *corner* of M is the angular sector between two consecutive edges around a vertex. A *rooted map* is a map with a marked corner, called the *root*; the vertex incident to the root is called the *root-vertex*. From now on, all maps are assumed to be rooted (note that the underlying graph of a rooted map is naturally vertex-rooted). A *unicellular map* is a map with a unique face. The classical Euler relation $|V| - |E| + |F| = 2 - 2g$ ensures that a unicellular map with n edges has $n + 1 - 2g$ vertices. A *plane tree* is a unicellular map of genus 0.

A *rotation system* on a connected graph G consists in a cyclic ordering of the half-edges of G around each vertex. Given a map M , its underlying graph G is naturally equipped with a rotation system given by the *clockwise ordering* of half-edges on the surface in a vicinity of each vertex. It is well-known that this correspondence is 1-to-1, i.e. a map can be considered as a connected graph equipped with a rotation system (thus, as a purely combinatorial object). We will take this viewpoint from now on.

A *cycle-signed permutation* is a permutation where each cycle carries a sign, either $+$ or $-$. A *C-permutation* is a cycle-signed permutation where all cycles have odd length, see Figure 1(a). For each C-permutation σ on n elements, the *rank* of σ is defined as $r(\sigma) = n - \ell(\sigma)$, where $\ell(\sigma)$ is the number of cycles of σ . Note that $r(\sigma)$ is even since all cycles have odd length. The *genus* of σ is defined as

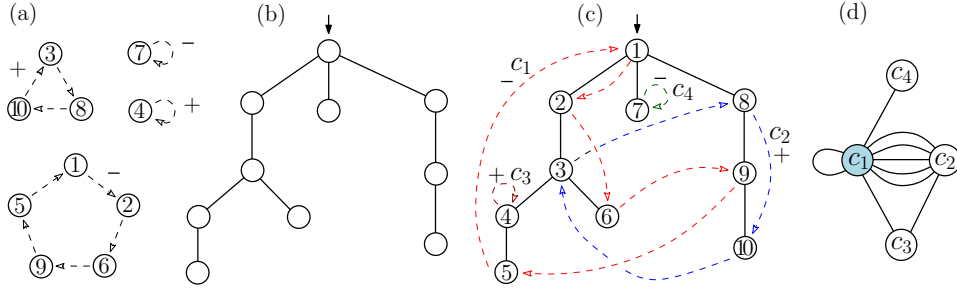


FIGURE 1. (a) A C-permutation σ . (b) A plane tree T . (c) The C-decorated tree (T, σ) . (d) The underlying graph of (T, σ) .

$r(\sigma)/2$. A *C-decorated tree* on n edges is a pair $\gamma = (T, \sigma)$ where T is a plane tree with n edges and σ is a C-permutation of $n+1$ elements. The *genus* of γ is defined to be the genus of σ . Note that the $n+1$ vertices of T can be canonically numbered from 1 to $n+1$ (e.g., following a left-to-right depth-first traversal), hence σ can be seen as a permutation of the vertices of T , see Figure 1(c). The *underlying graph* of γ is the (vertex-rooted) graph G obtained from T by merging into a single vertex the vertices in each cycle of σ (so that the vertices of G correspond to the cycles of σ), see Figure 1(d).

Definition 1. For n, g nonnegative integers, denote by $\mathcal{E}_g(n)$ the set of unicellular maps of genus g with n edges; and denote by $\mathcal{T}_g(n)$ the set of C-decorated trees of genus g with n edges.

For \mathcal{A} a finite set, $k\mathcal{A}$ denotes the set made of k disjoint copies of \mathcal{A} . For two finite sets \mathcal{A} and \mathcal{B} , we write $\mathcal{A} \simeq \mathcal{B}$ if there is a bijection between \mathcal{A} and \mathcal{B} . Our main result will be to show that $2^{n+1}\mathcal{E}_g(n) \simeq \mathcal{T}_g(n)$, with a bijection which preserves the underlying graphs of the objects.

2.2. Recursive decomposition of unicellular maps. In this section, we briefly recall a combinatorial method developed in [5] to decompose unicellular maps.

Proposition 1 (Chapuy [5]). For $k \geq 1$, denote by $\mathcal{E}_g^{(2k+1)}(n)$ the set of maps from $\mathcal{E}_g(n)$ in which a set of $2k+1$ vertices is distinguished. Then for $g > 0$ and $n \geq 0$,

$$(1) \quad 2g \mathcal{E}_g(n) \simeq \mathcal{E}_{g-1}^{(3)}(n) + \mathcal{E}_{g-2}^{(5)}(n) + \mathcal{E}_{g-3}^{(7)}(n) + \cdots + \mathcal{E}_0^{(2g+1)}(n).$$

In addition, if M and (M', S') are in correspondence, then the underlying graph of M is obtained from the underlying graph of M' by merging the vertices in S' into a single vertex.

We now sketch briefly the construction of [5]. Although this is not really needed for the sequel, we believe that it gives a good insight into the objects we are dealing with (readers in a hurry may take Proposition 1 for granted and jump directly to subsection 2.3). We refer to [5] for proofs and details.

We first explain where the factor $2g$ comes from in (1). Let M be a rooted unicellular map of genus g with n edges. Then M has $2n$ corners, and we label them from 1 to $2n$ incrementally, starting from the root, and going clockwise around the (unique) face of M (Figure 2). Let v be a vertex of M , let k be its degree, and

let (a_1, a_2, \dots, a_k) be the sequence of the labels of corners incident to it, read in *counterclockwise direction around v* starting from the minimal label $a_1 = \min a_i$. If for some $j \in \llbracket 1, k-1 \rrbracket$, we have $a_{j+1} < a_j$, we say that the corner of v labelled by a_{j+1} is a *trisection* of M . Figure 2(a) shows a map of genus two having four trisections. More generally we have:

Lemma 2 ([5]). *A unicellular map of genus g contains exactly $2g$ trisections. In other words, the set of unicellular maps of genus g with n edges and a marked trisection is isomorphic to $2g\mathcal{E}_g(n)$.*

Now, let τ be a trisection of M of label $\ell(\tau)$, and let v the vertex its belongs to. We denote by c the corner of v with minimum label and c' the corner with minimum label among those which appear between c and τ clockwise around v and whose label is greater than $\ell(\tau)$. By definition of a trisection, c' is well defined. We then construct a new map M' , by *slicing* the vertex v into three new vertices using the three corners c, c', τ as on Figure 2(b). We say that the map M' is obtained from M by *slicing the trisection τ* . As shown in [5], the new map M' is a unicellular map of genus $g-1$. We can thus relabel the $2n$ corners of M' from 1 to $2n$, according to the procedure we already used for M . Among these corners, three of them, say c_1, c_2, c_3 are naturally inherited from the slicing of v , as on Figure 2(b). Let v_1, v_2, v_3 be the vertices they belong to, respectively. Then the following is true [5]: *In the map M' , the corner c_i has the smallest label around the vertex v_i , for $i \in \{1, 2\}$. For $i = 3$, either the same is true, or c_3 is a trisection of the map M' .*

We now finally describe the bijection promised in Proposition 1. It is defined recursively on the genus, as follows. Given a map $M \in \mathcal{E}_g(n)$ with a marked trisection τ , let M' be obtained from M by the slicing of τ , and let c_i, v_i be defined as above for $i \in \{1, 2, 3\}$. If c_3 has the minimum label in v_3 , set $\Psi(M, \tau) := (M', \{v_1, v_2, v_3\})$, which is an element of $\mathcal{E}_{g-1}^{(3)}(n)$. Else, let $(M'', S) = \Psi(M', c_3)$, and set $\Psi(M, \tau) := (M'', S \cup \{v_1, v_2\})$. Note that this recursive algorithm necessarily stops, since the genus of the map decreases and since there are no trisections in unicellular maps of genus 0 (plane trees). Thus this procedure yields recursively a mapping that associates to a map M with a marked trisection τ another map M'' of a smaller genus, with a set S'' of marked vertices (namely the set of vertices which have been involved in a slicing at some point of the procedure). The set S'' of marked vertices necessarily has odd cardinality, as easily seen by induction. Moreover, it is clear that the underlying graph of M coincides with the underlying graph of M'' in which the vertices of S'' have been identified together into a single vertex. One can show that Ψ is a bijection by constructing explicitly the inverse mapping [5].

2.3. Recursive decomposition of C-decorated trees. We now propose a recursive method to decompose C-decorated trees, which can be seen as parallel to the decomposition of unicellular maps given in the previous section. Denote by $\mathcal{C}(n)$ (resp. $\mathcal{C}_g(n)$) the set of C-permutations on n elements (resp. on n elements and of genus g). A *signed sequence* of integers is a pair (ϵ, S) where S is an integer sequence and ϵ is a sign, either $+$ or $-$.

Lemma 3. *Let X be a finite non-empty set of positive integers. Then there is a bijection between signed sequences of distinct integers from X —all elements of X being present in the sequence— and C-permutations on the set X . In addition the*

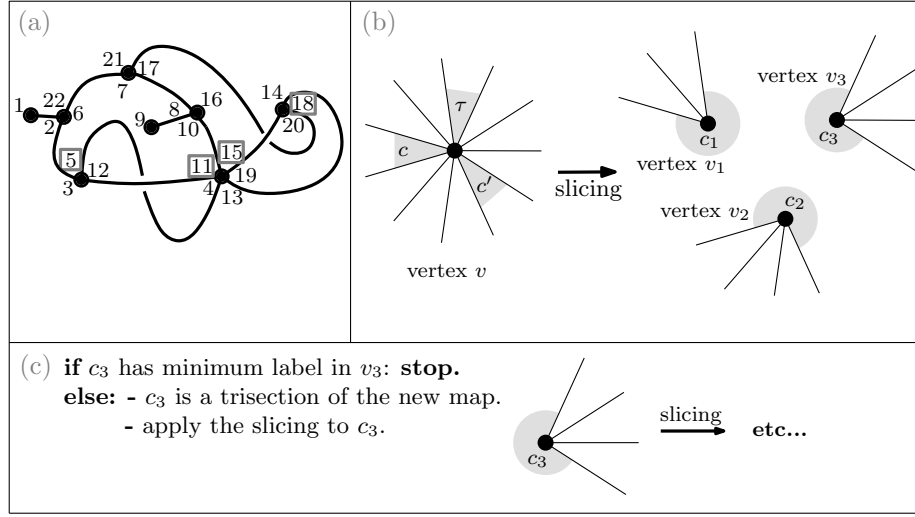


FIGURE 2. (a) A unicellular map of genus 2 equipped with its corner labelling. Labels corresponding to trisections are boxed. (b) Given a trisection τ , two other corners of interest c and c' are canonically defined (see text). “Slicing the trisection” then gives rise to three new vertices v_1, v_2, v_3 , with distinguished corners c_1, c_2, c_3 . (c) The recursive procedure of [5]: if c_3 is the minimum corner of v_3 , then stop; else, as shown in [5], c_3 is a trisection of the new map M' : in this case, iterate the slicing operation on (M', c_3) .

C-permutation has one cycle if and only if the signed sequence has odd length and starts with its minimal element.

Proof. Let γ be a signed sequence, e.g. $\gamma = {}^+(4731562)$. If γ has odd length and starts with its minimal element, return γ seen as a unicyclic C-permutation (where the unique cycle is written sequentially). Otherwise cut γ as $\gamma = \gamma_1\gamma_2$, where γ_2 starts with the minimal element in γ (in our example, $\gamma_1 = {}^+(473)$ and $\gamma_2 = (1562)$). If γ_2 has odd length, then “produce” the signed cycle ${}^+\gamma_2$. If γ_2 has even length, move the second element of γ_2 to the end of γ_1 , and “produce” the signed cycle ${}^-\gamma_2$. Then (in both cases), restart the same process on $\gamma = \gamma_1$, producing one (signed) cycle of odd length at each step, until γ is odd and starts with its minimal element, in which case one produces γ as the last signed cycle. (In our example, the signed cycles successively produced are ${}^-(162)$, ${}^-(3)$, and ${}^+(475)$.) The process clearly yields a collection of signed cycles of odd lengths, i.e., yields a C-permutation. The mapping is straightforward to invert, so it gives a bijection. \square

An element of a C-permutation is called *non-minimal* if it is not the minimum in its cycle. Non-minimal elements play the same role for C-permutations (and C-decorated trees) as trisections for unicellular maps. Indeed, a C-permutation of genus g has $2g$ non-minimal elements (compare with Lemma 2), and moreover we have the following analogue of Proposition 1:

Proposition 4. *For $k \geq 1$, denote by $\mathcal{T}_g^{(2k+1)}(n)$ the set of C -decorated trees from $\mathcal{T}_g(n)$ in which a set of $2k+1$ cycles is distinguished. Then for $g > 0$ and $n \geq 0$,*

$$2g \mathcal{T}_g(n) \simeq \mathcal{T}_{g-1}^{(3)} + \mathcal{T}_{g-2}^{(5)} + \mathcal{T}_{g-3}^{(7)} + \cdots.$$

In addition, if γ and (γ', S') are in correspondence, then the underlying graph of γ is obtained from the underlying graph of γ' by merging the vertices corresponding to cycles from S' into a single vertex.

Proof. For $k \geq 1$ let $\mathcal{C}_g^{(2k+1)}(n)$ be the set of C -permutations from $\mathcal{C}_g(n)$ where a subset of $2k+1$ cycles are marked. Let $\mathcal{C}_g^\circ(n)$ be the set of C -permutations from $\mathcal{C}_g(n)$ where a non-minimal element is marked. Note that $\mathcal{C}_g^\circ(n) \simeq 2g \mathcal{C}_g(n)$ since a C -permutation in $\mathcal{C}_g(n)$ has $2g$ non-minimal elements. Moreover $\mathcal{C}_g^\circ(n) \simeq \sum_{k \geq 1} \mathcal{C}_{g-k}^{(2k+1)}(n)$: apply Lemma 3 to the cycle —represented as a signed sequence— containing the marked non-minimal element, this produces a collection of $(2k+1) \geq 3$ signed cycles of odd length, which we take as the marked cycles. Hence $2g \mathcal{C}_g(n) \simeq \sum_{k \geq 1} \mathcal{C}_{g-k}^{(2k+1)}(n)$. Since $\mathcal{T}_g(n) = \mathcal{E}_0(n) \times \mathcal{C}_g(n+1)$, we conclude that $2g \mathcal{T}_g(n) \simeq \sum_{k \geq 1} \mathcal{T}_{g-k}^{(2k+1)}(n)$. The statement on the underlying graph just follows from the fact that the procedure in Lemma 3 merges the marked cycles into a unique cycle. \square

2.4. The main result.

Theorem 5. *For each non-negative integers n and g we have*

$$2^{n+1} \mathcal{E}_g(n) \simeq \mathcal{T}_g(n).$$

In addition the cycles of a C -decorated tree naturally correspond to the vertices of the associated unicellular map, in such a way that the respective underlying graphs are the same.

Proof. The proof is a simple induction on g , whereas n is fixed. The case $g = 0$ is obvious. Let $g > 0$. The induction hypothesis ensures that for each $g' < g$, $2^{n+1} \mathcal{E}_{g'}^{(2k+1)}(n) \simeq \mathcal{T}_{g'}^{(2k+1)}(n)$, where the underlying graphs (taking marked vertices into account) of corresponding objects are the same. Hence, by Propositions 1 and 4, we have $2g 2^{n+1} \mathcal{E}_g(n) \simeq 2g \mathcal{T}_g(n)$, where the underlying graphs of corresponding objects are the same. Finally, one can extract from this $2g$ -to- $2g$ correspondence a 1-to-1 correspondence (think of extracting a perfect matching from a $2g$ -regular bipartite graph, which is possible according to Hall's marriage theorem). And obviously the extracted 1-to-1 correspondence, which realizes $2^{n+1} \mathcal{E}_g(n) \simeq \mathcal{T}_g(n)$, also preserves the underlying graphs. \square

2.5. A fractional, or stochastic, formulation. Even if this does not hinder enumerative applications to be detailed in the next section, we do not know of an effective (polynomial-time) way to implement the bijection of Theorem 5; indeed the last step of the proof is to extract a perfect matching from a $2g$ -regular bipartite graph whose size is exponential in n .

What can be done effectively is a *fractional* formulation of the bijection. For a finite set X , let $\mathbb{C}\langle X \rangle$ be the set of linear combinations of the form $\sum_{x \in X} u_x \cdot x$, where the $x \in X$ are seen as independent formal vectors, and the coefficients u_x are in \mathbb{C} . Let $\mathbb{R}_1^+ \langle X \rangle \subset \mathbb{C}\langle X \rangle$ be the subset of linear combinations where the coefficients are nonnegative and add up to 1. Denote by $\mathbf{1}_X$ the vector $\sum_{x \in X} x$. For two finite

sets X and Y , a *fractional mapping* from X to Y is a linear mapping φ from $\mathbb{C}\langle X \rangle$ to $\mathbb{C}\langle Y \rangle$ such that the image of each $x \in X$ is in $\mathbb{R}_1^+\langle Y \rangle$; the subset of elements of Y with strictly positive coefficients in $\varphi(x)$ is called the *image-support* of x . Note that $\varphi(x)$ identifies to a probability distribution on Y ; a “call to $\varphi(x)$ ” is meant as picking up $y \in Y$ under this distribution. A fractional mapping is *bijective* if $\mathbf{1}_X$ is mapped to $\mathbf{1}_Y$, and is *deterministic* if each $x \in X$ is mapped to some $y \in Y$. Note that, if there is a fractional bijection from X to Y , then $|X| = |Y|$ (indeed in that case the matrix of φ is bistochastic).

One can now formulate by induction on the genus an effective (the cost of a call is $O(gn)$) fractional bijection from $2^{n+1}\mathcal{E}_g[n]$ to $\mathcal{T}_g(n)$, and similarly from $\mathcal{T}_g[n]$ to $2^{n+1}\mathcal{E}_g(n)$. The crucial property is that, for $k \geq 1$ and E, F finite sets, if there is a fractional bijection Φ from kE to kF then one can *effectively* derive from it a fractional bijection $\tilde{\Phi}$ from E to F : just define $\tilde{\Phi}(x)$ as $\frac{1}{k}(\iota(\Phi(x_1)) + \dots + \iota(\Phi(x_k)))$, where x_1, \dots, x_k are the representatives of x in kE , and where ι is the projection from kF to F . In other words a call to $\tilde{\Phi}(x)$ consists in picking up a representative x_i of x in kE uniformly at random and then calling $\Phi(x_i)$. Hence by induction on g , Propositions 1 and 4 (where the stated combinatorial isomorphisms are effective) ensure that there is an effective fractional bijection from $2^{n+1}\mathcal{E}_g(n)$ to $\mathcal{T}_g[n]$ and similarly from $\mathcal{T}_g[n]$ to $2^{n+1}\mathcal{E}_g[n]$, such that if γ' is in the image-support of γ then the underlying graphs of γ and γ' are the same.

Note that, given an effective fractional bijection between two sets X and Y , and a uniform random sampling algorithm on the set X , one obtains immediately a uniform random sampling algorithm for the set Y . In the next section, we will use our bijection to prove several enumerative formulas for unicellular maps, starting from elementary results on the enumeration of trees or permutations. In all cases, we will also be granted with a uniform random sampling algorithm for the corresponding unicellular maps, though we will not emphasize this point in the rest of the paper.

3. COUNTING FORMULAS FOR UNICELLULAR MAPS

It is quite clear that C-decorated trees are much simpler combinatorial objects than unicellular maps. In this section, we use them to give bijective proofs of several known formulas concerning unicellular maps. We focus on the Lehman-Walsh and Goupil-Schaeffer formulas, and the Harer-Zagier recurrence, of which bijective proofs were long-awaited. We also sketch a bijective proof of the Harer-Zagier summation formula (prototype for a family of formulas for which bijective proofs were already known). We insist on the fact that all these proofs are elementary consequences of our main bijection (Theorem 5).

3.1. Two immediate corollaries. The set $\mathcal{T}_g(n) = \mathcal{E}_0(n) \times \mathcal{C}_g(n+1)$ is the product of two sets that are easy to count. Precisely, let $\epsilon_g(n) = |\mathcal{E}_g(n)|$ and $c_g(n) = |\mathcal{C}_g(n)|$. Recall that $\epsilon_0(n) = \text{Cat}(n)$, where $\text{Cat}(n) := \frac{(2n)!}{n!(n+1)!}$ is the n th Catalan number. Therefore Theorem 5 gives $\epsilon_g(n) = 2^{-n-1}\text{Cat}(n)c_g(n+1)$.

It is immediate to give for $c_g(n+1)$ a closed form (by summing over all possible cycle types) or an explicit generating series. This yields two classical results for the enumeration of unicellular maps.

For $\gamma = (\gamma_1, \dots, \gamma_\ell) = 1^{m_1} \dots k^{m_k}$ a partition of g , the number $a_\gamma(n+1)$ of permutations of $n+1$ elements with cycle-type equal to $1^{n+1-2g-\ell} 3^{m_1} \dots (2k+1)^{m_k}$

is classically given by

$$a_\gamma(n+1) = \frac{(n+1)!}{(n+1-2g-\ell)! \prod_i m_i! (2i+1)^{m_i}},$$

and the number of C-permutations with this cycle-type is just $a_\gamma(n+1)2^{n+1-2g}$ (since each cycle has 2 possible signs). Hence, we get the equality $c_g(n+1) = 2^{n+1-2g} \sum_{\gamma \vdash g} a_\gamma(n+1)$. We thus recover:

Proposition 6 (Walsh and Lehman [21]). *The number $\epsilon_g(n)$ is given by*

$$\epsilon_g(n) = \frac{(2n)!}{n!(n+1-2g)!2^{2g}} \sum_{\gamma \vdash g} \frac{(n+1-2g)_\ell}{\prod_i m_i! (2i+1)^{m_i}},$$

where $(x)_k = \prod_{j=0}^{k-1} (x-j)$, ℓ is the number of parts of γ , and m_i is the number of parts of length i in γ .

The exponential generating function $C(x, y) := \sum_{n,g} \frac{1}{(n+1)!} c_g(n+1) y^{n+1} x^{n+1-2g}$ of C-permutations where y marks the number of elements, which are labelled, and x marks the number of cycles, is given by

$$C(x, y) = \exp\left(2x \sum_{k \geq 1} \frac{y^{2k+1}}{2k+1}\right) - 1 = \exp\left(x \log\left(\frac{1+y}{1-y}\right)\right) - 1 = \left(\frac{1+y}{1-y}\right)^x - 1.$$

Since $c_0(1) = 2$ and $\frac{1}{(n+1)!} c_g(n+1) = \frac{2^{n+1} n!}{(2n)!} \epsilon_g(n) = \frac{2}{(2n-1)!!} \epsilon_g(n)$ for $n \geq 1$, we recover:

Proposition 7 (Harer-Zagier series formula [10, 12]). *The generating function*

$$E(x, y) := 1 + 2xy + 2 \sum_{g \geq 0, n > 0} \frac{\epsilon_g(n)}{(2n-1)!!} y^{n+1} x^{n+1-2g} \text{ is given by}$$

$$E(x, y) = \left(\frac{1+y}{1-y}\right)^x.$$

3.2. Harer-Zagier recurrence formula. Elementary algebraic manipulations on the expression of $E(x, y)$ yield a very simple recurrence satisfied by $\epsilon_g(n)$, known as the Harer-Zagier recurrence formula (stated in Proposition 10 hereafter). We now show that the model of C-decorated trees makes it possible to derive this recurrence directly from a combinatorial isomorphism, that generalizes Rémy's beautiful bijection for plane trees [17].

It is convenient here to consider C-decorated trees as *unlabelled structures*: precisely we see a C-decorated tree as a plane tree where the vertices are partitioned into parts of odd size, where each part carries a sign $+$ or $-$, and such that the vertices in each part are cyclically ordered (the C-permutation can be recovered by numbering the vertices of the tree according to a left-to-right depth-first traversal), think of Figure 1(c) where the labels have been taken out. We take here the convention that a plane tree with n edges has $2n+1$ corners, considering that the sector of the root has two corners, one on each side of the root.

We denote by $\mathcal{P}(n) = \mathcal{E}_0(n)$ the set of plane trees with n edges, and by $\mathcal{P}^v(n)$ (resp. $\mathcal{P}^c(n)$) the set of plane trees with n edges where a vertex (resp. a corner) is marked. Rémy's procedure, shown in Figure 3, realizes the isomorphism $\mathcal{P}^v(n) \simeq 2\mathcal{P}^c(n-1)$, or equivalently

$$(2) \quad (n+1)\mathcal{P}(n) \simeq 2(2n-1)\mathcal{P}(n-1).$$

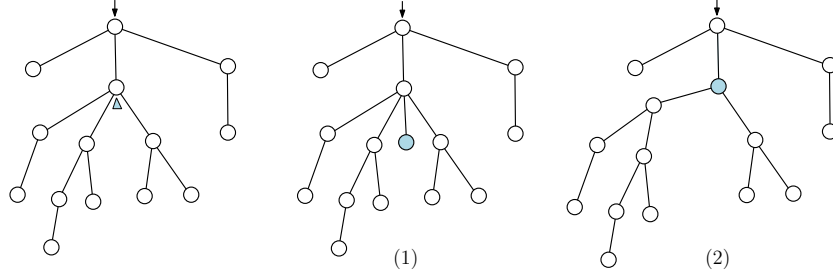


FIGURE 3. Rémy's procedure gives two ways to obtain a plane tree with n edges and a marked vertex v from a plane tree with $n-1$ edges and a marked corner: (1) in the first way (replacing the marked corner by a leg) v is a leaf, (2) in the second way (stretching an edge to carry the subtree on the left of the marked corner) v is a non-leaf.

Let $\mathcal{T}_g^v(n)$ be the set of C-decorated trees from $\mathcal{T}_g(n)$ where a vertex is marked. Let \mathcal{A} (resp. \mathcal{B}) be the subset of objects in $\mathcal{T}_g^v(n)$ where the signed cycle containing the marked vertex has length 1 (resp. length greater than 1). Let $\gamma \in \mathcal{T}_g^v(n)$, with $n \geq 1$. If $\gamma \in \mathcal{A}$, record the sign of the 1-cycle containing v and then apply the Rémy's procedure to the plane tree with respect to v (so as to delete v). This reduction, which does not change the genus, yields $\mathcal{A} \simeq 2 \cdot 2(2n-1)\mathcal{T}_g(n-1)$. If $\gamma \in \mathcal{B}$, let c be the cycle containing the marked vertex v ; c is of the form $(v, v_1, v_2, \dots, v_{2k})$ for some $k \geq 1$. Move v_1 and v_2 out of c (the successor of v becomes the former successor of v_2). Then apply the Rémy's procedure twice, firstly with respect to v_1 (on a plane tree with n edges), secondly with respect to v_2 (on a plane tree with $n-1$ edges). This reduction, which decreases the genus by 1, yields $\mathcal{B} \simeq 2(2n-1)2(2n-3)\mathcal{T}_{g-1}^v(n-2)$, hence $\mathcal{B} \simeq 4(n-1)(2n-1)(2n-3)\mathcal{T}_{g-1}(n-2)$. Since $\mathcal{T}_g^v(n) = \mathcal{A} + \mathcal{B}$ and $\mathcal{T}_g^v(n) \simeq (n+1)\mathcal{T}_g(n)$, we finally obtain the isomorphism

$$(3) \quad (n+1)\mathcal{T}_g(n) \simeq 4(2n-1)\mathcal{T}_g(n-1) + 4(n-1)(2n-1)(2n-3)\mathcal{T}_{g-1}(n-2),$$

which holds for any $n \geq 1$ and $g \geq 0$ (with the convention $\mathcal{T}_g(n) = \emptyset$ if g or n is negative). Since $2^{n+1}\mathcal{E}_g(n) \simeq \mathcal{T}_g(n)$, we recover:

Proposition 8 (Harer-Zagier recurrence formula [10, 12]). *The coefficients $\epsilon_g(n)$ satisfy the following recurrence relation valid for any $g \geq 0$ and $n \geq 1$ (with $\epsilon_0(0) = 1$ and $\epsilon_g(n) = 0$ if $g < 0$ or $n < 0$):*

$$(n+1)\epsilon_g(n) = 2(2n-1)\epsilon_g(n-1) + (n-1)(2n-1)(2n-3)\epsilon_{g-1}(n-2).$$

To the best of our knowledge this is the first proof of the Harer-Zagier recurrence formula that directly follows from a combinatorial isomorphism. The isomorphism (3) also provides a natural extension to arbitrary genus of Rémy's isomorphism (2).

3.3. Refined enumeration of bipartite unicellular maps. In this paragraph, we explain how to recover a formula due to Goupil and Schaeffer [9, Theorem 2.1] from our bijection. Let us first give a few definitions. A graph is *bipartite* if its vertices can be colored in black and white such that each edge connects a black and

a white vertex. If the graph has a root-vertex v , then v is required to be black; and if the graph is also connected, then such a bicolouration of the vertices is unique. From now on, a connected bipartite graph with a root-vertex is assumed to be endowed with this canonical bicolouration.

The degree distribution of a map/graph is the sequence of the degrees of its vertices taken in decreasing order (it is a partition of $2n$, where n is the number of edges). If we consider a bipartite map/graph, we can consider separately the *white vertex degree distribution* and the *black vertex degree distribution*, which are two partitions of n .

Let ℓ, m, n be positive integers such that $n + 1 - \ell - m$ is even. Fix two partitions λ, μ of n of respective lengths ℓ and m . We call $\text{Bi}(\lambda, \mu)$ the number of bipartite unicellular maps, with white (resp. black) vertex degree distribution λ (resp. μ). The corresponding genus is $g = (n + 1 - \ell - m)/2$. It will be convenient to change a little bit the formulation of the problem and to consider *labelled maps* instead of the usual non-labelled maps: a *labelled map* is a map whose vertices are labelled with integers $1, 2, \dots$. If the map is bipartite, we require instead that the white and black vertices are labelled separately (with respective labels w_1, w_2, \dots and b_1, b_2, \dots). The degree distribution(s) of a (bipartite) labelled map with n edges can be seen as a composition of $2n$ (resp. two compositions of n). For $\mathbf{I} = (i_1, \dots, i_\ell)$ and $\mathbf{J} = (j_1, \dots, j_m)$ two compositions of n , we denote by $\text{BiL}(\mathbf{I}, \mathbf{J})$ the number of labelled bipartite unicellular maps with white (resp. black) vertex degree distribution \mathbf{I} (resp. \mathbf{J}). The link between $\text{Bi}(\lambda, \mu)$ and $\text{BiL}(\mathbf{I}, \mathbf{J})$ is straightforward: $\text{BiL}(\mathbf{I}, \mathbf{J}) = m_1(\lambda)!m_2(\lambda)! \cdots m_1(\mu)!m_2(\mu)! \cdots \text{Bi}(\lambda, \mu)$, where λ and μ are the sorted versions of \mathbf{I} and \mathbf{J} . We now recover the following formula:

Proposition 9 (Goupil and Schaeffer [9, Theorem 2.1]).

$$(4) \quad \text{BiL}(\mathbf{I}, \mathbf{J}) = 2^{-2g} \cdot n \cdot (\ell + 2g_1 - 1)!(m + 2g_2 - 1)! \\ \cdot \sum_{g_1 + g_2 = g} \sum_{\substack{p_1 + \dots + p_\ell = g_1 \\ q_1 + \dots + q_m = g_2}} \prod_{r=1}^{\ell} \frac{1}{2p_r + 1} \binom{i_r - 1}{2p_r} \prod_{r=1}^m \frac{1}{2q_r + 1} \binom{j_r - 1}{2q_r}.$$

Proof. For $g = 0$ the formula is simply

$$(5) \quad \text{BiL}(\mathbf{I}, \mathbf{J}) = n(\ell - 1)!(m - 1)!,$$

which can easily be established by a bivariate version of the cyclic lemma, see also [7, Theorem 2.2]. (Note, that in that case, the cardinality only depends on the lengths of \mathbf{I} and \mathbf{J} .)

We now prove the formula for arbitrary g . Consider some lists $\mathbf{p} = (p_1, \dots, p_\ell)$ and $\mathbf{q} = (q_1, \dots, q_m)$ of nonnegative integers with total sum g : let $g_1 = \sum p_i$ and $g_2 = \sum q_i$. We say that a composition \mathbf{H} refines \mathbf{I} along \mathbf{p} if \mathbf{H} is of the form $(h_1^1, \dots, h_1^{2p_1+1}, \dots, h_\ell^1, \dots, h_\ell^{2p_\ell+1})$, with $\sum_{t=1}^{2p_r+1} h_r^t = i_r$ for all r between 1 and ℓ . Clearly, there are $\prod_{r=1}^{\ell} \binom{i_r - 1}{2p_r}$ such compositions \mathbf{H} . One defines similarly a composition \mathbf{K} refining \mathbf{J} along \mathbf{q} .

Consider now the set of labelled bipartite plane trees of vertex degree distributions \mathbf{H} and \mathbf{K} , where \mathbf{H} (resp. \mathbf{K}) refines \mathbf{I} (resp. \mathbf{J}) along \mathbf{p} (resp. \mathbf{q}). By (5), there are $n \cdot (\ell + 2g_1 - 1)!(m + 2g_2 - 1)!$ trees for each pair (\mathbf{H}, \mathbf{K}) , so in total, with

\mathbf{I} , \mathbf{J} , \mathbf{p} and \mathbf{q} fixed, the number of such trees is:

$$(6) \quad n \cdot (\ell + 2g_1 - 1)!(m + 2g_2 - 1)! \prod_{r=1}^{\ell} \binom{i_r - 1}{2p_r} \prod_{r=1}^m \binom{j_r - 1}{2q_r}.$$

As the parts of \mathbf{H} (resp. \mathbf{K}) are naturally indexed by pairs of integers, we can see these trees as labelled by the set $\{w_r^t; 1 \leq r \leq \ell, 1 \leq t \leq 2p_r + 1\} \sqcup \{b_r^t; 1 \leq r \leq m, 1 \leq t \leq 2q_r + 1\}$. There is a canonical permutation of the vertices of the trees with cycles of odd sizes and which preserves the bicoloration: just send w_r^t to w_r^{t+1} (resp. b_r^t to b_r^{t+1}), where $t + 1$ is meant modulo $2p_r + 1$ (resp. $2q_r + 1$). If we additionally put a sign on each cycle, we get a C-decorated tree (with labelled cycles) that corresponds to a labelled bipartite map with white (resp. black) vertex degree distribution \mathbf{I} (resp. \mathbf{J}). Conversely, to recover a labelled bipartite plane tree from such a C-decorated tree, one has to choose in each cycle which vertex gets the label w_r^1 or b_r^1 , and one has to forget the signs of the $(n + 1 - g)$ cycles. This represents a factor $2^{n+1-2g} \left(\prod_{r=1}^{\ell} (2p_r + 1) \prod_{r=1}^m (2q_r + 1) \right)^{-1}$.

Multiplying (6) by the above factor, and summing over all possible sequences \mathbf{p} and \mathbf{q} of total sum g , we conclude that the number of C-decorated trees associated with labelled bipartite unicellular maps of white (resp. black) vertex degree distribution \mathbf{I} (resp. \mathbf{J}), is equal to 2^{n+1} times the right-hand side of (4). By Theorem 5, this number is also equal to $2^{n+1} \text{BiL}(\mathbf{I}, \mathbf{J})$. This ends the proof of Proposition 9. \square

This is the first combinatorial proof of (4) (the proof by Goupil and Schaeffer involves representation theory of the symmetric group). Moreover, the authors of [9] found surprising that “the two partitions contribute independently to the genus”. With our approach, this is very natural, since the cycles are carried independently by white and black vertices.

3.4. Counting colored maps. In this paragraph, we deal with what was presented in the introduction as the *first family* of formulas. These formulas give an expression for a certain *sum* of counting coefficients of unicellular maps, the expressions being usually simpler than those for single counting coefficients of unicellular maps (like the Goupil-Schaeffer’s formula). These sums can typically be seen as counting formulas for *colored unicellular maps* (where the control is on the number of colors, which gives indirect access to the genus).

3.4.1. A summation formula for unicellular maps. We begin by Harer-Zagier’s summation formula [10, 12] (which can also be very easily derived from the expression of $E(x, y)$). In contrast to the formulas presented so far, this one has already been given combinatorial proofs [13, 8, 1] using different bijective constructions, but we want to insist on the fact that our construction gives bijective proofs for all the formulas in a unified way.

Proposition 10 (Harer-Zagier summation formula [10, 12]). *Let $A(v; n)$ be the number of unicellular maps with n edges and v vertices. Then for $n \geq 1$*

$$\sum_v A(v; n) x^v = (2n - 1)!! \sum_{r \geq 1} 2^{r-1} \binom{n}{r-1} \binom{x}{r}.$$

Proof. It suffices to prove that the number $A_r(n)$ of unicellular maps with n edges, each vertex having a color in $[1..r]$, and each color in $[1..r]$ being used at least once, is given by $A_r(n) = (2n-1)!! 2^{r-1} \binom{n}{r-1}$. Consider a C-decorated tree with n edges, where each (signed) cycle has a color in $[1..r]$, and such that each color in $[1..r]$ is used by at least one cycle. Each of the r colors yields a (non-empty) C-permutation, which can be represented as a signed sequence, according to Lemma 3. Then one can concatenate these r signed sequences into a unique sequence S of length $n+1$, together with r signs and a subset of $r-1$ elements among the n elements from position 2 to $n+1$ in S (in order to recover from S the r signed sequences). For instance if $r=3$ and if the signed sequences corresponding respectively to colors 1, 2, 3 are $^+(3, 9, 4)$, $^-(5, 8, 6, 2)$, and $^-(1, 7)$, then the concatenated sequence is $(3, 9, 4, 5, 8, 6, 2, 1, 7)$, together with the 3 signs $(+, -, -)$ and the two selected elements $\{5, 1\}$. Hence the number of such C-decorated trees is $(n+1)! 2^r \binom{n}{r-1}$, and by Theorem 5,

$$A_r(n) = 2^{-n-1} \text{Cat}(n) (n+1)! 2^r \binom{n}{r-1} = (2n-1)!! 2^{r-1} \binom{n}{r-1}. \quad \square$$

3.4.2. A summation formula for bipartite unicellular maps. By Theorem 5, a C-decorated tree associated to a bipartite unicellular map is a bipartite plane tree such that each signed cycle must contain only white (resp. black) vertices. Recall that the $n+1$ vertices carry distinct labels from 1 to $n+1$ (the ordering follows by convention a left-to-right depth-first traversal, see Figure 1(c)). Without loss of information the i black vertices (resp. j white vertices) can be relabelled from 1 to i (resp. from 1 to j) in the order-preserving way; we take here this convention for labelling the vertices of such a C-decorated tree. We now recover the following summation formula due to Jackson (different bijective proofs have been given in [18] and in [1]):

Proposition 11 (Jackson's summation formula [11]). *Let $B(v, w; n)$ be the number of bipartite unicellular maps with n edges, v black vertices and w white vertices. Then for $n \geq 1$*

$$\sum_{v,w} B(v, w; n) y^v z^w = n! \sum_{r,s \geq 1} \binom{n-1}{r-1, s-1} \binom{y}{r} \binom{z}{s}.$$

Proof. It suffices to prove that, for $r, s \geq 1$, the number $B_{r,s}(n)$ of bipartite unicellular maps with n edges, each black (resp. white) vertex having a so-called *b-color* in $[1..r]$ (resp. a so-called *w-color* in $[1..s]$), such that each b-color in $[1..r]$ (resp. w-color in $[1..s]$) is used at least once, is given by $B_{r,s}(n) = n! \binom{n-1}{r-1, s-1}$. For n, i, j such that $i+j = n+1$, consider a bipartite C-decorated tree with n edges, i black vertices, j white vertices, where each black (resp. white) signed cycle has a b-color in $[1..r]$ (resp. a w-color in $[1..s]$), and each b-color in $[1..r]$ (resp. w-color in $[1..s]$) is used at least once. By the same argument as in Proposition 10, the C-permutation and b-colors on black vertices can be encoded by a sequence S_b of length i of distinct integers in $[1..i]$, together with r signs and a subset of $r-1$ elements among the $i-1$ elements at positions from 2 to i in S_b . And the C-permutation and w-colors on white vertices can be encoded by a sequence S_w of length j of distinct integers in $[1..j]$, together with s signs and a subset of $s-1$ elements among the $j-1$ elements at positions from 2 to j in S_w . Hence there are $\text{Nar}(i, j; n) 2^{r+s} i! \binom{i-1}{r-1} j! \binom{j-1}{s-1}$ such

C-decorated trees, where $\text{Nar}(i, j; n)$ (called the *Narayana* number) is the number of bipartite plane trees with n edges, i black vertices and j white vertices, given by $\text{Nar}(i, j; n) = \frac{1}{n} \binom{n}{i} \binom{n}{j}$. By Theorem 5,

$$\begin{aligned} B_{r,s}(n) &= 2^{-n-1} 2^{r+s} \sum_{i+j=n+1} \text{Nar}(i, j; n) i! j! \binom{i-1}{r-1} \binom{j-1}{s-1} \\ &= n!(n-1)! \frac{2^{r+s-n-1}}{(r-1)!(s-1)!} \sum_{i+j=n+1} \frac{1}{(i-r)!(j-s)!}. \end{aligned}$$

Next we have

$$\sum_{i+j=n+1} \frac{1}{(i-r)!(j-s)!} = \sum_{i+j=n+1-r-s} \frac{1}{i!j!} = \frac{2^{n+1-r-s}}{(n+1-r-s)!}.$$

Hence $B_{r,s}(n) = n! \binom{n-1}{r-1, s-1}$. \square

3.4.3. A refinement. A. Morales and E. Vassilieva [14] have established a very elegant summation formula for bipartite maps, counted with respect to their degree distributions, which can be viewed as a refinement of Jackson's summation formula (indeed, it is an easy exercise to recover Jackson's summation formula out of it):

Proposition 12 (Morales and Vassilieva [14, Theorem 1]). *Let m_λ and p_ρ be the monomial and power sum basis of the ring of symmetric functions and \mathbf{x} and \mathbf{y} two infinite sets of variables. Then, for any $n \geq 1$,*

$$\sum_{\lambda, \mu \vdash n} \text{Bi}(\lambda, \mu) p_\lambda(\mathbf{x}) p_\mu(\mathbf{y}) = \sum_{\rho, \nu \vdash n} \frac{n(n - \ell(\rho))!(n - \ell(\nu))!}{(n+1 - \ell(\rho) - \ell(\nu))!} m_\rho(\mathbf{x}) m_\nu(\mathbf{y}).$$

The original proof given in [14] goes through a complicated bijection with newly introduced objects called *thorn trees* by the authors. The bijective method in [1] (which is well adapted to summation formulas) also makes it possible to get the formula. And a short non-bijective proof has been given recently in [20] using characters of the symmetric groups and Schur functions. We explain here how this result can be recovered from our bijection. The proof is very similar to the one of Goupil-Schaeffer's formula.

Let us first recall that Proposition 12 can be reformulated in purely combinatorial terms (without symmetric functions) using colored maps.

By definition here, a bipartite unicellular map is *colored* by associating to each white (resp. black) vertex a color in $[1 \dots \ell_w]$ (resp. $[1 \dots \ell_b]$), each color between 1 and ℓ_w (resp. ℓ_b) being chosen at least once (note: we always think of the color r of a white vertex as *different* from the color r of a black vertex). To a colored bipartite map with n edges one can associate its *colored degree distribution*, that is, the pair (\mathbf{I}, \mathbf{J}) of compositions of n , where the k th part of \mathbf{I} (resp. of \mathbf{J}) is the sum of the degrees of the white (resp. black) vertices with color k .

We denote by $\text{BiC}(\mathbf{I}, \mathbf{J})$ the number of colored bipartite unicellular maps of colored degree distribution (\mathbf{I}, \mathbf{J}) . Then Proposition 12 is equivalent to the following statement [14, paragraph 2.4]:

Proposition 13. *For any compositions \mathbf{I} and \mathbf{J} of the same integer n which satisfy $\ell(\mathbf{I}) + \ell(\mathbf{J}) \leq n + 1$,*

$$\text{BiC}(\mathbf{I}, \mathbf{J}) = \frac{n(n - \ell(\mathbf{I}))!(n - \ell(\mathbf{J}))!}{(n + 1 - \ell(\mathbf{I}) - \ell(\mathbf{J}))!}.$$

Proof. Let us consider two compositions \mathbf{H} and \mathbf{K} which refine respectively \mathbf{I} and \mathbf{J} , and such that $\ell(\mathbf{H}) + \ell(\mathbf{K}) = n + 1$. A part of \mathbf{H} (resp. of \mathbf{K}) is said to have color r if it is contained in the r th part of \mathbf{I} (resp. of \mathbf{J}).

Consider a labelled bipartite tree T whose white vertex degrees (in the order given by the labels) follow the composition \mathbf{H} and whose black vertex degrees follow the composition \mathbf{K} . A black (resp. white) vertex is said to have color r if the corresponding part of \mathbf{H} (resp. of \mathbf{K}) has color r . Since the tree is labelled, the white (resp. black) vertices with the same color r are totally ordered as (w_r^1, w_r^2, \dots) (resp. (b_r^1, b_r^2, \dots)). Hence if we add the data of a sign per color ($2^{\ell(\mathbf{I}) + \ell(\mathbf{J})}$ choices for all signs), using Lemma 3, we can see the vertices with the same color as endowed with a C -permutation.

Putting all these C -permutations together, we obtain a C -permutation of the vertices of the tree T , which has the following property: the vertices in the same cycle always have the same color. Applying our main bijection (Theorem 5), we obtain a bipartite unicellular map. The vertices of this map have a canonical coloration, as each vertex corresponds to a cycle of the C -permutation. By construction, this colored map has colored degree distribution (\mathbf{I}, \mathbf{J}) .

To sum up, by Theorem 5 and the construction above, each colored bipartite unicellular map with colored degree distribution (\mathbf{I}, \mathbf{J}) can be obtained in 2^{n+1} different ways from

- a labelled bipartite tree T of white (resp. black) vertex degree given by \mathbf{H} (resp. \mathbf{K}) for *some* refinements \mathbf{H} and \mathbf{K} with $\ell(\mathbf{H}) + \ell(\mathbf{K}) = n + 1$;
- the assignment of a sign to each color.

The number of possible signs is always $2^{\ell(\mathbf{I}) + \ell(\mathbf{J})}$, so this yields a constant factor. For given compositions \mathbf{H} and \mathbf{K} , the number of corresponding trees is

$$n(\ell(\mathbf{H}) - 1)!(\ell(\mathbf{K}) - 1)!$$

Thus we have to count the number of refinements \mathbf{H} (resp. \mathbf{K}) of \mathbf{I} (resp. \mathbf{J}) with a given value ℓ of $\ell(\mathbf{H})$ (resp. m of $\ell(\mathbf{K})$). It is easily seen to be equal to

$$\binom{n - \ell(\mathbf{I})}{\ell - \ell(\mathbf{I})} \text{ (resp. } \binom{n - \ell(\mathbf{J})}{m - \ell(\mathbf{J})} \text{)}.$$

Finally, by Theorem 5, we get:

$$2^{n+1} \text{BiC}(\mathbf{I}, \mathbf{J}) = 2^{\ell(\mathbf{I}) + \ell(\mathbf{J})} \sum_{\substack{\ell + m = n + 1 \\ \ell \geq \ell(\mathbf{I}), m \geq \ell(\mathbf{J})}} n(\ell - 1)!(m - 1)! \binom{n - \ell(\mathbf{I})}{\ell - \ell(\mathbf{I})} \binom{n - \ell(\mathbf{J})}{m - \ell(\mathbf{J})}.$$

Denoting $h = n + 1 - \ell(\mathbf{I}) - \ell(\mathbf{J})$ and setting $h_1 = \ell - \ell(\mathbf{I})$, $h_2 = m - \ell(\mathbf{J})$ in the summation index, the right hand side of the previous equation writes as:

$$2^{\ell(\mathbf{I}) + \ell(\mathbf{J})} \sum_{h_1 + h_2 = h} n(\ell(\mathbf{I}) + h_1 - 1)!(\ell(\mathbf{J}) + h_2 - 1)! \binom{n - \ell(\mathbf{I})}{h_1} \binom{n - \ell(\mathbf{J})}{h_2}.$$

But the relation $\ell(\mathbf{I}) + \ell(\mathbf{J}) + h_1 + h_2 = n + 1$ implies that $(\ell(\mathbf{J}) + h_2 - 1)! \binom{n - \ell(\mathbf{I})}{h_1} = \frac{(n - \ell(\mathbf{I}))!}{h_1!}$ and $(\ell(\mathbf{I}) + h_1 - 1)! \binom{n - \ell(\mathbf{J})}{h_2} = \frac{(n - \ell(\mathbf{J}))!}{h_2!}$. Plugging this in the expression

above, we get

$$\begin{aligned} 2^{n+1}\text{BiC}(\mathbf{I}, \mathbf{J}) &= 2^{\ell(\mathbf{I})+\ell(\mathbf{J})} \cdot n \cdot (n - \ell(\mathbf{I}))! \cdot (n - \ell(\mathbf{J}))! \sum_{h_1+h_2=h} \frac{1}{h_1!h_2!} \\ &= 2^{\ell(\mathbf{I})+\ell(\mathbf{J})} \cdot n \cdot (n - \ell(\mathbf{I}))! \cdot (n - \ell(\mathbf{J}))! \frac{2^h}{h!}. \end{aligned}$$

The powers of 2 cancel each other and we get the desired result. \square

3.5. Covered maps, shuffles, and an identity of [3]. *Covered maps* were introduced in [3] as an extension of the notion of tree-rooted map (map equipped with a spanning tree). A covered map of genus g is a rooted map M of genus g , non necessarily unicellular, equipped with a distinguished connected subgraph S (with the same vertex set as M) having the following property:

viewed as a map, S is a unicellular map, possibly of a different genus than M .

Here, in order to view S “as a map”, we equip it with the map structure induced by M : the clockwise ordering of half-edges of S around each vertex is defined as the restriction of the clockwise ordering in M (see [3] for details). The genus g_1 of S is an element of $\llbracket 0, g \rrbracket$. For example, $g_1 = 0$ if and only if S is a spanning tree of M . In general, we say that the covered map (M, S) has *type* (g, g_1) .

Covered maps have an interesting duality property that generalizes the existence of dual spanning trees in the planar case: namely, each covered map (M, S) of type (g, g_1) has a dual covered map (M^*, S') of type (g, g_2) with $g_1 + g_2 = g$. By extending ideas of Mullin [15], it is not difficult to describe the covered map M as a “shuffle” of the two unicellular maps S and S' , see [3]. It follows that the number $\text{Cov}_{g_1, g_2}(n)$ of covered maps of type $(g_1 + g_2, g_1)$ with n edges can be expressed as the following shuffle-sum [3, eq. (6)]:

$$(7) \quad \text{Cov}_{g_1, g_2}(n) = \sum_{n_1+n_2=n} \binom{2n}{2n_1} \epsilon_{g_1}(n_1) \epsilon_{g_2}(n_2).$$

In the case $g_1 = g_2 = 0$, this sum simplifies thanks to the Chu-Vandermonde identity, and we have the remarkable result due to Mullin [15] (see [2] for a bijective proof):

$$(8) \quad \text{Cov}_{0,0}(n) = \text{Cat}(n) \text{Cat}(n+1).$$

The main enumerative result of the paper [3] is a generalisation of (8) to any genus, obtained via a difficult bijection:

Proposition 14 (Bernardi and Chapuy, [3]). *For all $n \geq 1$ and $g \geq 0$, the number $\text{Cov}_g(n) = \sum_{g_1+g_2=g} \text{Cov}_{g_1, g_2}(n)$ of covered maps of genus g with n edges is equal to:*

$$\text{Cov}_g(n) = \text{Cat}(n) \text{Bip}_g(n+1),$$

where $\text{Bip}_g(n+1)$ is the number of rooted bipartite unicellular maps of genus g with $n+1$ edges. Equivalently, the following identity holds:

$$(9) \quad \sum_{g_1+g_2=g} \sum_{n_1+n_2=n} \binom{2n}{2n_1} \epsilon_{g_1}(n_1) \epsilon_{g_2}(n_2) = \text{Cat}(n) \text{Bip}_g(n+1).$$

Proof. We denote as before by $c_g(m)$ the number of C -permutations of genus g of a set of m elements. By our main result, Theorem 5, the left-hand side of (9) can be rewritten as:

$$2^{-n-2} \sum_{g_1+g_2=g} \sum_{n_1+n_2=n} \binom{2n}{2n_1} c_{g_1}(n_1+1) c_{g_2}(n_2+1) \text{Cat}(n_1) \text{Cat}(n_2).$$

We now observe that:

$$\binom{2n}{2n_1} \text{Cat}(n_1) \text{Cat}(n_2) = \text{Cat}(n) \text{Nar}(n_1+1, n_2+1; n+1)$$

where as before the Narayana number $\text{Nar}(i, j; n)$ is the number of bipartite plane trees with n edges, i black vertices and j white vertices (this last equality follows directly from the explicit expressions of Catalan and Narayana numbers; an interpretation in terms of planar tree-rooted maps is given by the bijection of [2]). Therefore we have:

$$\text{Cov}_g(n) = 2^{-n-2} \text{Cat}(n) \sum_{g_1+g_2=g} \sum_{n_1+n_2=n} c_{g_1}(n_1+1) c_{g_2}(n_2+1) \text{Nar}(n_1+1, n_2+1; n+1).$$

Now, the double-sum in this equation is equal to the number of bipartite C -decorated trees (that is, bipartite trees equipped with a C -permutation of the vertices that stabilizes each color class) with $n+1$ edges and genus g : indeed in the double-sum, the quantities g_1 and n_1+1 can be interpreted respectively as the genus of the restriction of the C -permutation to black vertices of the tree, and as the number of black vertices in the tree. By our main result, Theorem 5, this double-sum is therefore equal to $2^{n+2} \text{Bip}_g(n+1)$, which proves (9). \square

The proof above and (7) also show the following fact. Let G_1 be the genus of the submap S in a covered map (M, S) of genus g with n edges chosen uniformly at random, and let G_\circ be the genus of the restriction to white vertices of the C -permutation in a bipartite C -decorated tree of genus g with $n+1$ edges chosen uniformly at random. Then the random variables G_1 and G_\circ have the same distribution.

It is possible to prove that, when g is fixed and n tends to infinity, the variable G_\circ is close to a binomial random variable $B(g, 1/2)$: the idea behind this property is that a random bipartite tree with $n+1$ edges has about $n/2 + O(\sqrt{n})$ vertices of each color with high probability, and that with high probability the C -permutation of its vertices is made of g cycles of length 3, that independently “fall” into each of the color classes with probability $1/2$. Giving a proper proof of these elementary statements would lead us to far from our main subject, so we leave to the reader the details of a proof along these lines of the following fact, which was proved in [3] with no combinatorial interpretation:

Proposition 15 ([3]). *Let $g \geq g_1 \geq 0$. When n tends to infinity, the probability that a covered map of genus g with n edges chosen uniformly at random has type (g, g_1) tends to $2^{-g} \binom{g}{g_1}$.*

To conclude this section, we mention that, in [3], refined results were given that take more parameters into account (e.g., the number of vertices and faces of the covered map). These extensions can be proved exactly in the same way as Proposition 14, but we do not state them explicitly here, for the sake of brevity.

4. COMPUTING STANLEY CHARACTER POLYNOMIALS

4.1. Formulation of the problem. We now consider the following enumerative problem. For n a fixed integer, we would like to compute the generating series

$$F_n(p_1, p_2, \dots; q_1, q_2, \dots) = \sum_{(M, \varphi)} \text{wt}(M, \varphi)$$

of pairs (M, φ) where M is a rooted bipartite unicellular map with n edges, and φ is a mapping from the vertex set V_M of M to positive integers, satisfying the following *order condition*:

for each edge e of M , one has $\varphi(b_e) \geq \varphi(w_e)$, where b_e and w_e are respectively the black and white extremities of e .

The weight of such a pair is $\text{wt}(M, \varphi) := \prod_{v \in V_M^\circ} p_{\varphi(i)} \prod_{v \in V_M^\bullet} q_{\varphi(i)}$, where V_M^\bullet and V_M° are respectively the sets of black (resp. white) vertices of M .

Our motivation comes from representation theory of the symmetric group. This topic is linked to map enumeration by the following formula conjectured in [19] and proved in [6]. Let $\mathbf{p} = p_1, \dots, p_r$ and $\mathbf{q} = q_1, \dots, q_r$ be two finite lists of positive integers of the same length. Then the evaluation of the generating series considered above is equal to

$$(10) \quad F_n(p_1, \dots, p_r, 0, \dots; q_1, \dots, q_r, 0, \dots) = L(L-1) \cdots (L-n+1) \hat{\chi}^\lambda((1 \ 2 \ \cdots \ n)),$$

where:

- λ is the partition with p_1 parts equal to $q_1 + \dots + q_r$, p_2 parts equal to $q_2 + \dots + q_r$, and so on...
- $L = \sum_{1 \leq i \leq j \leq r} p_i q_j$ is its number of boxes ;
- $\hat{\chi}^\lambda$ is the normalized character of the irreducible representation of S_L associated to λ ;
- $(1 \ 2 \ \cdots \ n)$ is an n -th cycle seen as a permutation of S_L (if $n > L$, it is not defined but, as the numerical factor is 0, it is not a problem).

Remark 1. In [19, 6], this formula is stated under a slightly different form. We call G_n the same generating series as F_n except that the order condition is replaced by the following *maximum condition*:

for each black vertex b , one has $\varphi(b) = \max \varphi(w)$, where the maximum is taken over all white neighbours w of b .

Then the main theorem of [6] states that

$$G_n(p'_1, \dots, p'_r, 0, \dots; q'_1, \dots, q'_r, 0, \dots) = L(L-1) \cdots (L-n+1) \hat{\chi}^\lambda((1 \ 2 \ \cdots \ n)),$$

where everything is defined as above except that

λ is the partition with p'_1 parts equal to q'_1 , p'_2 parts equal to q'_2 , and so on...

This result is clearly equivalent to (10) by setting:

$$\forall i \geq 1, \begin{cases} p_i = p'_i \\ q_i = q'_i - q'_{i+1} \end{cases}.$$

4.2. A new expression for F_n . Our main bijection allows us to express the generating series F_n in terms of the corresponding generating series for plane trees:

$$R_{n+1}(\mathbf{p}, \mathbf{q}) = \sum_{(T, \varphi)} \text{wt}(T, \varphi),$$

where the sum runs over all pairs (T, φ) , T being a plane tree and φ a function $V_T \rightarrow \mathbb{N}$ satisfying the order condition.

The strange notation R_{n+1} comes from the following fact: A. Rattan has proved [16] that this generating series is the $n + 1$ -th free cumulant R_{n+1} of the transition measure of the Young diagram λ (λ states here for the Young diagram defined in terms on \mathbf{p} and \mathbf{q} in the previous paragraph). Free cumulants have become in the last few years an important tool in (asymptotic) representation theory of the symmetric groups, see for example the work of P. Biane [4].

Let us define an operator D by

$$D(x^k) := \sum_{g \geq 0} c_g(k) x^{k-2g} = k! \sum_{r=1}^k 2^r \binom{k-1}{r-1} \binom{x}{r},$$

D being extended multiplicatively to monomials in distinct variables, and then extended linearly to multivariate polynomials and series (in particular, series in the variables \mathbf{p} and \mathbf{q}).

Theorem 16. *For any $n \geq 1$, one has $2^{n+1}F_n = D(R_{n+1})$.*

Proof. A pair (M, φ) as above corresponds by the bijection of Theorem 5 to a bipartite C -decorated tree T , together with a function $\varphi : V_T \rightarrow \mathbb{N}$ which fulfills the order condition and such that all vertices in a given cycle have the same image by φ . The result follows directly. \square

The free cumulant R_{n+1} is the compositional inverse of an explicit series [16]. Hence Theorem 16 gives an efficient, easily implemented way of computing Stanley character polynomials F_n .

REFERENCES

- [1] O. Bernardi. An analogue of the Harer-Zagier formula for unicellular maps on general surfaces. arXiv:1011.2311, 2010.
- [2] Olivier Bernardi. Bijective counting of tree-rooted maps and shuffles of parenthesis systems. *Electron. J. Combin.*, 14(1):Research Paper 9, 36 pp. (electronic), 2007.
- [3] Olivier Bernardi and Guillaume Chapuy. A bijection for covered maps, or a shortcut between Harer-Zagier's and Jackson's formulas. *J. Combin. Theory Ser. A*, 118(6):1718–1748, 2011.
- [4] P. Biane. Representations of symmetric groups and free probability. *Adv. Math.*, 138(1):126–181, 1998.
- [5] G. Chapuy. A new combinatorial identity for unicellular maps, via a direct bijective approach. *Advances in Applied Mathematics*, 47(4):874 – 893, 2011.
- [6] V. Féray. Stanley's formula for characters of the symmetric group. *Ann. Comb.*, 13(4):453–461, 2010.
- [7] I. P. Goulden and D. M. Jackson. The combinatorial relationship between trees, cacti and certain connection coefficients for the symmetric group. *European J. Combin.*, 13(5):357–365, 1992.
- [8] I. P. Goulden and A. Nica. A direct bijection for the Harer-Zagier formula. *J. Comb. Theory, Ser. A*, 111(2):224–238, 2005.
- [9] A. Goupil and G. Schaeffer. Factoring n -cycles and counting maps of given genus. *European J. Combin.*, 19(7):819–834, 1998.

- [10] J. Harer and D. Zagier. The Euler characteristic of the moduli space of curves. *Invent. Math.*, 85:457–486, 1986.
- [11] D. M. Jackson. Some combinatorial problems associated with products of conjugacy classes of the symmetric group. *J. Comb. Theory, Ser. A*, 49(2):363–369, 1988.
- [12] S. K. Lando and A. K. Zvonkin. *Graphs on Surfaces and Their Applications*. Springer, 2004.
- [13] B. Lass. Démonstration combinatoire de la formule de Harer-Zagier. *C. R. Acad. Sci. Paris*, 333, Série I:155–160, 2001.
- [14] A. Morales and E. Vassilieva. Bijective enumeration of bicolored maps of given vertex degree distribution. *DMTCS Proceedings*, AK:661–672, 2009.
- [15] R. C. Mullin. On the enumeration of tree-rooted maps. *Canad. J. Math.*, 19:174–183, 1967.
- [16] A. Rattan. Stanley’s character polynomials and coloured factorizations in the symmetric group. *J. Combin. Theory Ser. A*, 114(4):535–546, 2008.
- [17] J.-L. Rémy. Un procédé itératif de dénombrement d’arbres binaires et son application à leur génération aléatoire. *RAIRO Inform. Théor.*, 19(2):179–195, 1985.
- [18] G. Schaeffer and E. A. Vassilieva. A bijective proof of Jackson’s formula for the number of factorizations of a cycle. *J. Comb. Theory, Ser. A*, 115(6):903–924, 2008.
- [19] R. P. Stanley. A conjectured combinatorial interpretation of the normalized irreducible character values of the symmetric group. *arxiv:math/0606467*, 2006.
- [20] Ekaterina A Vassilieva. Explicit monomial expansions of the generating series for connection coefficients. *arXiv:1111.6215*, 2011.
- [21] T. R. S. Walsh and A. B. Lehman. Counting rooted maps by genus. I. *J. Combin. Theory Ser. B*, 13:192–218, 1972.

(G.C.) CNRS, LIAFA - UMR 7089, UNIVERSITÉ PARIS 7, 75205 PARIS CEDEX, FRANCE.

(V.F.) CNRS, LABRI - UMR 5800, UNIVERSITÉ DE BORDEAUX, 33405 TALENCE CEDEX, FRANCE.

(É.F.) CNRS, LIX - UMR 7161, ÉCOLE POLYTECHNIQUE, 91128 PALAISEAU CEDEX, FRANCE.